SURVIVAL ANALYSIS OF

AIDS DRUG ASSISTANCE PROGRAM (ADAP) CLIENTS

FEBRUARY 2005

ARNOLD SCHWARZENEGGER
Governor
State of California

Kimberly Belshé
Secretary
Health and Human Services Agency

Sandra Shewry
Director
Department of Health Services

SURVIVAL ANALYSIS OF

AIDS DRUG ASSISTANCE PROGRAM (ADAP) CLIENTS

Prepared by:

Dennis T. Wong, Ph.D.
Biao Xing, M.A.

Department of Health Services
Office of AIDS
HIV/AIDS Epidemiology Branch
http://www.dhs.ca.gov/AIDS

Kevin Reilly, D.V.M., M.P.V.M.
Deputy Director
Prevention Services

Michael Montgomery, Chief
Office of AIDS

Juan Ruiz, M.D., Dr.P.H., M.P.H., Chief
HIV/AIDS Epidemiology Branch
Office of AIDS

February 2005

ACKNOWLEDGMENTS

I would like to thank Susan M. Sabatier and Kathleen Russell for providing valuable feedback on this report.

Correspondence
Please send any questions or comments to Dennis T. Wong: **dwong2@dhs.ca.gov**.

Suggested Citation
Wong, D.T., Xing, B., *Survival Analysis of AIDS Drug Assistance Program (ADAP) Clients*. California Department of Health Services, Office of AIDS, 2005.

TABLE OF CONTENTS

TABLES

TABLE OF CONTENTS (Continued)

FIGURES

EXECUTIVE SUMMARY

Objectives. The Department of Health Services, Office of AIDS (DHS/OA) attempted to answer the following research questions: 1) How long do ADAP clients live?; 2) How long do AIDS-diagnosed ADAP clients live?; 3) What are the percentages of AIDS-diagnosed clients when broken out by mode of exposure or risk category?; and 4) Do factors such as gender, race/ethnicity, age group, geographical location, and other variables affect the survival rates for either group?

Design. ADAP clients enrolled and served between January 1, 1998, and December 31, 2000, were matched with death certificate records for California residents who died between 1998-2001 (ADAP-to-Vital Statistics match) and all California AIDS cases as of July 31, 2002, in the HIV/AIDS Reporting System (HARS) (ADAP-to-HARS match). Survival analyses were performed with respect to time to death since: 1) HIV diagnosis; and 2) AIDS diagnosis.

Results. Due to the low case fatality rate (cases/deaths), median survival times could not be reliably estimated. In the overall HIV/AIDS ADAP population, African Americans and Hispanic/Latinos had a higher rate of death since HIV diagnosis, or shorter survival times since HIV diagnosis, than Whites. In the AIDS-diagnosed ADAP population, in contrast, Whites and Hispanic/Latinos had a higher rate of death since AIDS diagnosis than African Americans. In both HIV/AIDS- and AIDS-diagnosed populations, older age groups tended to have a higher rate of death than younger age groups. Also, clients who received all of their HIV-related prescriptions through ADAP and those with more prescription access months had a smaller rate of death than those with Medicaid and/or private insurance benefits and those with fewer access months.

Conclusions. Race/ethnicity differences in ADAP survival rates since HIV diagnosis suggest that people of color may delay seeking treatment and that people of color may encounter socioeconomic and cultural barriers to educational and preventive services, primary care and treatment, and access to costly highly active antiretroviral therapy (HAART) drug therapies. Differences among the AIDS-diagnosed ADAP population were not conclusive because of the low number of deaths in ADAP. Regardless, ADAP serves a diverse population and should continue monitoring the demographics of its clients. Overall, HIV/AIDS health care providers, practitioners, and advocates should recognize that it may be more difficult for African Americans and Hispanic/Latinos to access HIV care and treatment. Through efforts to provide the best HIV/AIDS services for the different socioeconomic levels and cultural values of the population served, barriers to HIV health care can be prevented.

INTRODUCTION

Survival analysis examines the time interval between two events, the entry point into a study and the occurrence of a particular event. For example, how long do AIDS-diagnosed individuals live? However, the observation period may differ for all participants, and the terminal event may not happen for everyone. Survival analysis uses special statistical techniques to account for these and other factors.

ADAP

The California ADAP was enacted legislatively in October 1987. The program, funded by the state and federal Ryan White Comprehensive AIDS Resources Emergency Act, provides HIV-related therapies to low- to middle-income uninsured or under-insured persons living with HIV/AIDS (PLWH/A). In July 1996, protease inhibitors were added to the ADAP formulary following Food and Drug Administration approval. Shortly thereafter, ADAP clients were able to take advantage of triple combination therapies known as HAART. As the use of these promising but costly new therapies became more widespread and client enrollment in ADAP increased, state and federal appropriations for the program increased dramatically to allow for program expansion.

ADAP continues to monitor monthly client demographics and program drug expenditures. However, limited client information collected by ADAP precludes a more in-depth evaluation of the impact ADAP has on the health outcomes of the population it serves. A recent study by OA integrated death and hospital records containing HIV/AIDS-related information with the ADAP database over a three-year period (1997-99), which allowed us to better understand the role ADAP played in the quality of life for the individuals it served (Wong and Xing, 2003).

The results found race/ethnicity and age differences in the number of deaths and hospitalizations among ADAP clients. For example, Whites had higher mortality rates than African Americans and Hispanic/Latinos in 1998. The following year, Hispanic/Latinos had lower mortality rates than both groups. In 1998 and 1999, older age groups had higher mortality rates, more hospital visits, and longer lengths of stay than their younger counterparts. Also, clients on “preferred” HAART had lower mortality rates and slightly fewer hospital visits than those on “non-preferred” drug combinations (as defined by the federal Guidelines for the Use of Antiretroviral Agents in HIV-Infected Adults and Adolescents by the Panel on Clinical Practices for Treatment of HIV Infection, May 5, 1999).

To estimate the probability of death in ADAP, a logistic regression model of deaths was built, which yielded similar results to the above findings; that is, Hispanic/Latino deaths were less likely than Whites, older clients were more likely to die than younger ones, and those accessing more preferred treatment regimens based on federal guidelines were less likely to die than those accessing less amounts of preferred regimens.

While the study answered some important questions, it also raised others. For example, how long do ADAP clients live? Is this timeframe comparable to that of "typical" AIDS-diagnosed individuals? Are there any demographic differences in ADAP survival rates? At the national level, the Centers for Disease Control and Prevention (CDC) reported that of 394,705 Americans diagnosed with AIDS between 1984-1997, median survival time increased from 11 months for 1984 diagnoses to 46 months for 1995 diagnoses (Lee, Karon, Selik, Neal, and Fleming, 2001). At the state level, median survival time for California AIDS cases has increased from 10 months for 1985 diagnoses to 34 months for 1993 diagnoses (Bryan and Sun, 1998). In 1994, California median survival rates improved dramatically to nearly 80 months, which may reflect, in part, fewer deaths and more variability in the limited amount of data (OA, 2001). A more recent study found that private insurance was more effective than public insurance in reducing HIV/AIDS mortality (Bhattacharya, Goldman, and Sood, 2003).

THE STUDY

This study was an extension to our previous research effort described above. OA continued to examine the effectiveness of ADAP by determining the survival times of all PLWH/A clients and its AIDS-diagnosed subpopulation. Specifically, OA attempted to answer the following research questions:

1. How long do ADAP clients live (i.e., survival time)?

2. How long do AIDS-diagnosed ADAP clients live?

3. What are the percentages of AIDS-diagnosed clients when broken out by mode of exposure or risk category?

4. Do factors such as gender, race/ethnicity, age group, geographical location, and other variables affect the survival rates for either group?

Since no previous study has examined survival data in ADAP, no specific hypotheses were formulated for this study.

METHOD AND PROCEDURES

Data Sets

ADAP. ADAP enrolled and served 30,309 (unduplicated) PLWH/A clients between January 1, 1998, and December 31, 2000. It was estimated that 40 percent of these clients were AIDS diagnosed based on annual report data.

Demographic variables of interest included gender (male or female), race/ethnicity (White, African American, Hispanic/Latino, or Other), age group (ages 18-30, 31-40, 41-50, or over 50), and geographic location (Title I/urban areas or Title II/rural areas).

Additional variables examined included AIDS diagnosis (yes or no), insurance coverage group (ADAP Only or Not ADAP Only—ADAP with Medicaid share-of-cost or private insurance co-payments), and access months (the number of months a client accessed a drug prescription through ADAP).

Death Statistical Master Files. Death certificate records for California residents who died between 1998-2001 were obtained through DHS Office of Vital Records, Vital Statistics Section. During the four-year period, January 1, 1998, and December 31, 2001, 926,945 deaths were reported.

HARS. OA maintains HARS, a computer database containing demographic and clinical information on all California AIDS cases. As of July 31, 2002, a cumulative total of 126,269 AIDS cases had been reported in California. Of these cases, 76,529 had died; a case fatality rate of 61 percent.

Demographic variables of interest were the same as above for ADAP. Additional variables collected from HARS were: 1) AIDS diagnosis date; 2) Mode of exposure (MSM—men who have sex with men, IDU—injection drug use, MSM and IDU, heterosexual contact, and Other—e.g., hemophilia/coagulation disorder or risk not reported/Other); and 3) Date of death.

Matching Process

Both ADAP and Vital Statistics death files contain names and Social Security Numbers (SSNs) in its records. HARS, in contrast, primarily uses a "soundex" code to facilitate matching and unduplicating AIDS cases. The soundex code maintains an individual's anonymity by converting the individual's last name into a four-digit alphanumeric code. The first letter of his/her last name becomes an index letter followed by a three-digit number. Thirteen general rules are applied to create the final soundex code (see Appendix A). Because HARS is a non-name reporting system, names and SSNs are usually not available until a person's death. All three data sets included demographic variables.

To link ADAP with Vital Statistics death files and HARS, OA used a probabilistic matching process based on potential common identifiers (soundex, names, SSN, gender, race/ethnicity, and date of birth (DOB) (see Appendix B). The first match, ADAP-to-Vital Stats, yielded 2,320 clients who had died (eight percent). The second match, ADAP-to-HARS, resulted in 13,224 AIDS-diagnosed clients out of 30,309 (44 percent).

Data Quality Management. As a quality check of the ADAP-to-HARS match, OA examined the most recent HIV/AIDS diagnosis of these 13,224 clients from the ADAP data set to see how many clients were identified in HARS as AIDS diagnosed. Excluding 1,952 unknown cases, 14 percent were HIV asymptomatic, 29 percent were HIV symptomatic, and 57 percent were AIDS diagnosed. This was a much larger percentage of HIV-diagnosed clients than expected. Reasons for the discrepancy include different levels of completeness in the matching identifiers (i.e., missing data), a lack of diagnosis updates in ADAP, and reporting delays and data entry errors in both ADAP and HARS.

RESULTS

ADAP Population

Demographics. Table 1 shows the demographic characteristics of the ADAP PLWH/A population in 1998-2000. Of the 30,309 clients, 89 percent were males and 10 percent were females. In the race/ethnicity category, 45 percent were White, 17 percent were African American, 30 percent were Hispanic/Latino, and the remaining seven percent were Other/Unknown. Fifteen percent of ADAP population were between 18-30 years old, 45 percent were 31-40 years old, 29 percent were 41-50 years old, and 10 percent were over 50 years old. Ninety-two percent of clients resided in Title I/urban areas and seven percent resided in Title II/rural areas.

TABLE 1: DEMOGRAPHICS FOR ADAP POPULATION

GENDER	FREQ	PCT
Male	27,132	89.30%
Female	3,166	10.42%
Other/Unknown	11	0.28%
RACE/ETHNICITY	**FREQ**	**PCT**
White	13,667	45.02%
African American	5,245	17.29%
Hispanic/Latino	9,220	30.29%
Other/Unknown	2,177	7.40%
AGE GROUP	**FREQ**	**PCT**
Under 18 years old	12	0.04%
18-30 years old	4,562	15.03%
31-40 years old	13,777	45.34%
41-50 years old	8,812	29.02%
Over 50 years old	3,106	10.45%
Unknown	40	0.13%
GEOGRAPHICAL LOCATION	**FREQ**	**PCT**
Title I/Urban Areas	27,960	92.03%
Title II/Rural Areas	2,044	6.73%
Unknown	305	1.24%
TOTAL	**30,309**	**100.00%**

Table 2 shows the frequencies and percentages on other variables for the ADAP population. Forty-three percent were AIDS diagnosed via the ADAP-to-HARS match, while the remaining 56 percent were HIV diagnosed only. Among insurance coverage group, 67 percent were ADAP Only and 33 percent had ADAP with Medicaid or private insurance benefits. Mean access months, or number of months a client accessed a prescription in three years, was 13.75 with a standard deviation of 11.15.

TABLE 2: DESCRIPTIVE STATISTICS ON OTHER VARIABLES FOR ADAP POPULATION		
AIDS DIAGNOSIS	**FREQ**	**PCT**
Yes	13,171	43.46%
No	17,065	56.30%
Unknown	73	0.24%
COVERAGE GROUP	**FREQ**	**PCT**
ADAP Only	20,212	66.69%
Not ADAP Only	10,097	33.31%
TOTAL	**30,309**	**100.00%**
For Coverage Group, Not ADAP Only includes ADAP with Medicaid or private insurance benefits.		

ADAP-to-Vital Statistics Survival Analysis

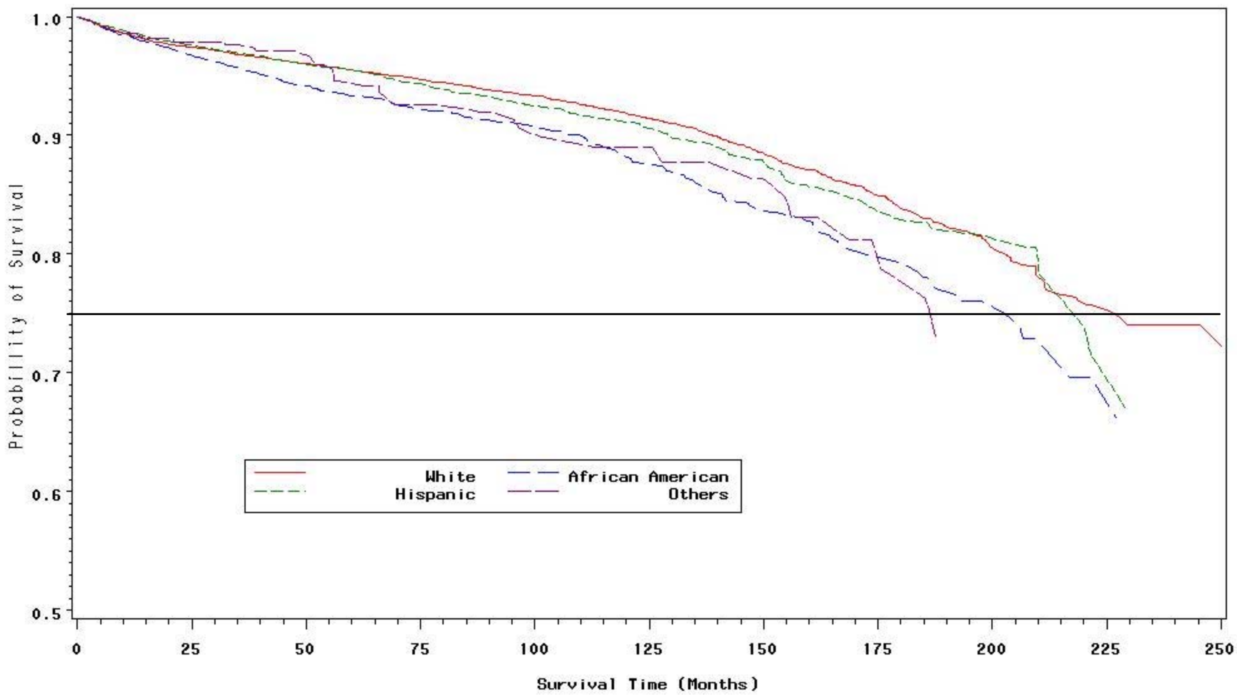
Survival Curves. OA examined the Kaplan-Meier estimate of the 25 percentile of survival time since HIV diagnosis, or time in months when 25 percent of the clients were expected to die.[1] Because of the small number of deaths, the 50 percentile (median) and the 75 percentile of survival times could not be reliably estimated. Figure 1 shows the survival curves by race/ethnicity. Whites had the longest 25 percentile of survival time (227.84 months), followed by Hispanic/Latinos (220.03 months), and African Americans (202.56 months). The other/unknown category had the shortest survival estimates (187.57 months). The log-rank test statistic ($\underline{p}$<.000) suggested that there were significant differences in survival times across different race/ethnicity groups, without adjusting other variables.

Figure 2 shows the survival patterns by age group. Longer 25 percentile of survival time estimates were found among middle-aged clients (31-40 = 227.84 months and 41-50 = 221.48 months) than their younger (18-30 = 205.84 months) and older (Over 50 = 190.56 months) counterparts. The log-rank test statistic ($\underline{p}$<.000) suggested that there were significant differences in survival times across different age groups, without adjusting other variables.

Among AIDS diagnosis, HIV clients (273.18 months) had longer 25 percentile of survival time than AIDS-diagnosed clients (164.72 months), and among insurance coverage group, ADAP Only clients (249.97 months) had longer 25 percentile of survival time than Not ADAP Only clients (204.131 months). These results are illustrated in Figures 3 and 4, respectively. In addition, the log-rank test statistics corresponding to the figures indicated significant differences between groups ($\underline{p}$<.000).

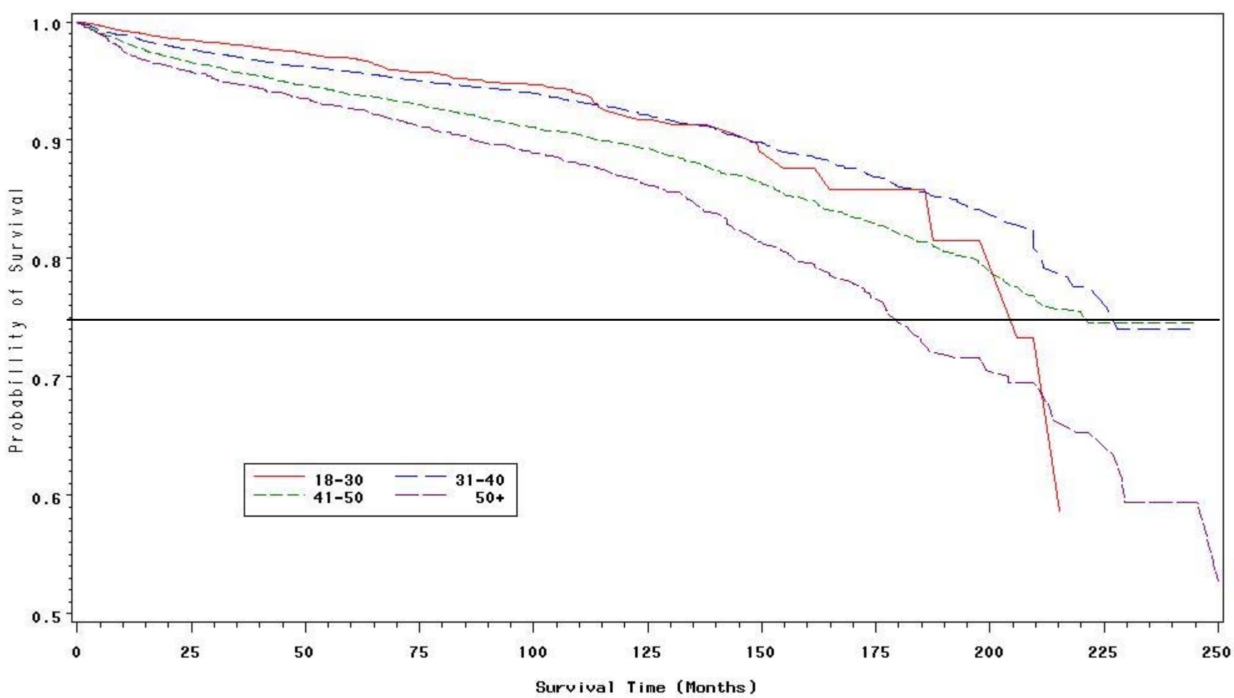
[1] Only survival curves for categorical variables with a significant log-rank test are shown.

Figure 1. Kaplan-Meier Estimate of Survival Curves by Race/Ethnicity



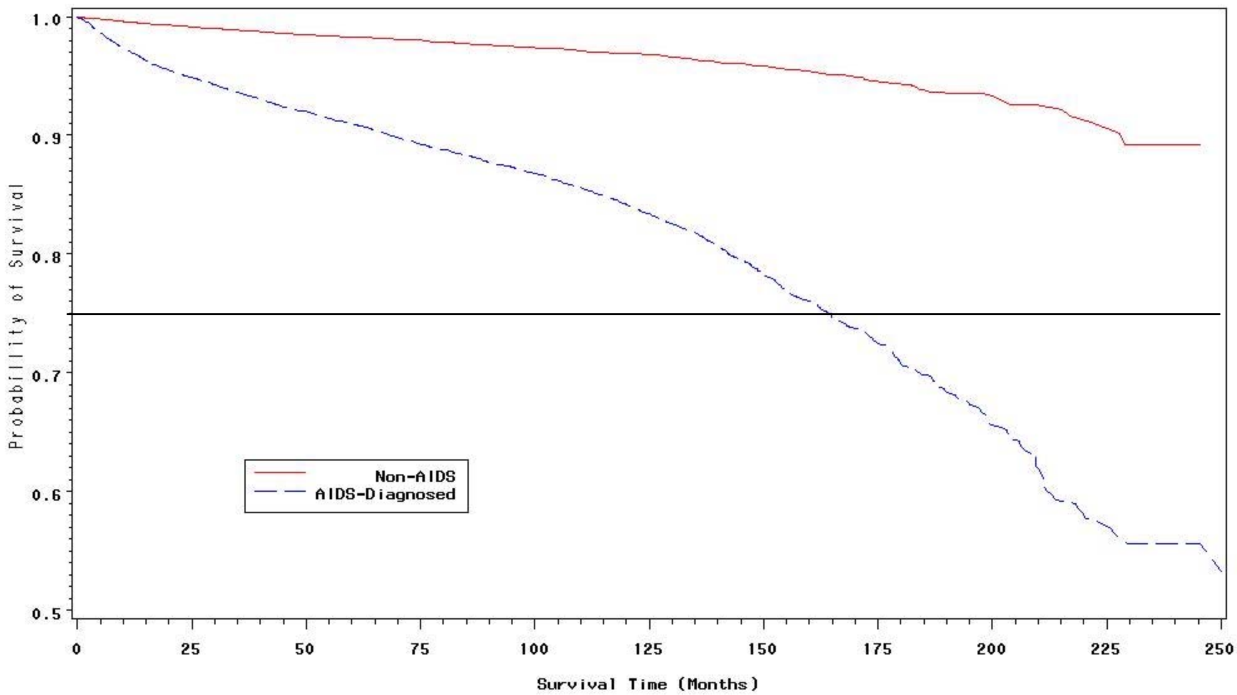
Department of Health Services, Office of AIDS, AIDS Drug Assistance Program

Figure 2. Kaplan-Meier Estimate of Survival Curves by Age Group



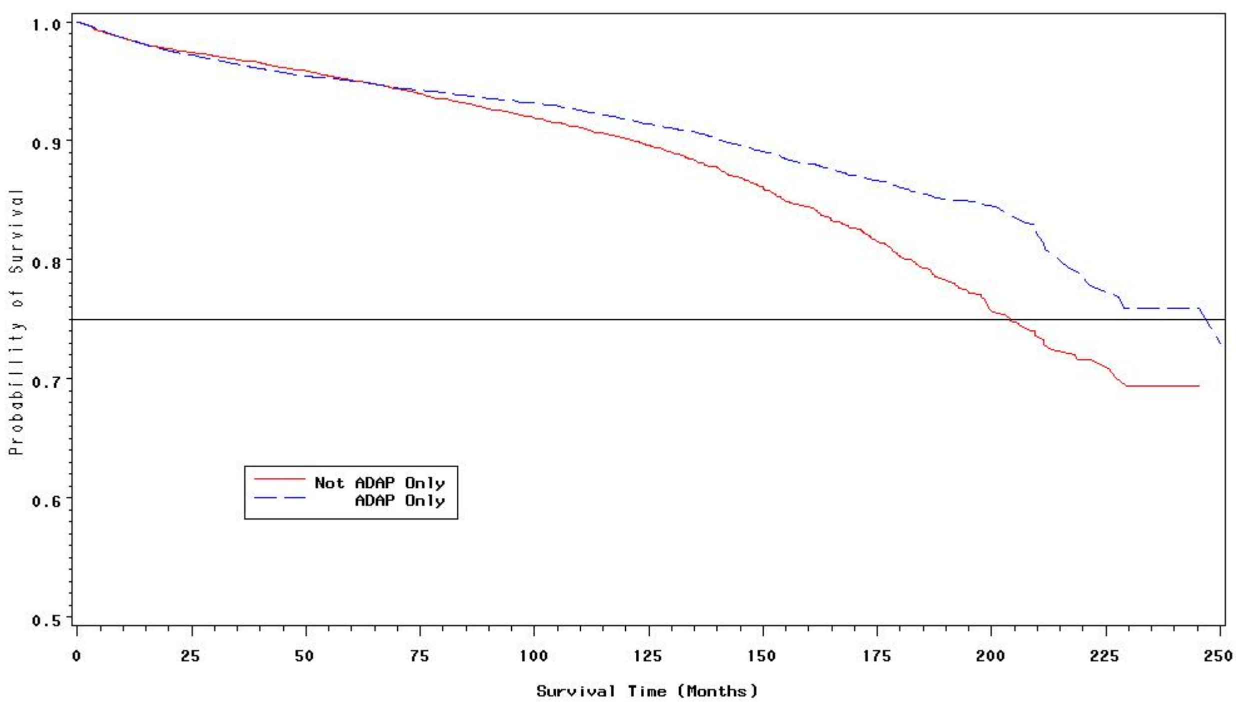
Department of Health Services, Office of AIDS, AIDS Drug Assistance Program

Figure 3. Kaplan-Meier Estimate of Survival Curves by AIDS Diagnosis



Department of Health Services, Office of AIDS, AIDS Drug Assistance Program

Figure 4. Kaplan-Meier Estimate of Survival Curves by Insurance Coverage Group



Department of Health Services, Office of AIDS, AIDS Drug Assistance Program

<u>Time to Death Since HIV Diagnosis</u>. To estimate the adjusted effect of variables of interest on survival time since HIV diagnosis, a Cox-Proportional Hazard Model was used with demographic and other variables included. Table 3 shows these results. Clients without a HIV diagnosis date were deleted from this analysis leaving a total of 28,184 cases including 2,160 deaths.

Significant differences in death hazard ratios (HRs) were found on race/ethnicity, age group, AIDS diagnosis, insurance coverage group, and access months. The HR (95 percent Confidence Intervals [CI]) was higher for African Americans (HR = 1.13, CI = 1.00, 1.26) and Hispanics/Latinos (HR = 1.20, CI = 1.08, 1.34) than Whites, indicating Whites had longer survival time than people of color. Significantly higher HRs were found within the age groups 41-50 (HR = 1.57, CI = 1.31, 1.87) and Over 50 (HR = 2.26, CI = 1.86, 2.74) than 18-30. The largest HR was 5.40 (CI = 4.83, 6.03) on AIDS diagnosis. Smaller HRs were found for ADAP Only clients in comparison to Not ADAP Only clients (HR = 0.83, CI = 0.75, 0.90), and clients with more access months (HR = 0.94 for one more access months, CI = 0.93, 0.94). These results indicated that clients who were ADAP Only and those with more access months had longer survival time than their counterparts.

Additional analyses were performed on all pairwise comparisons for race/ethnicity and age group. No differences were found between other race/ethnicity groups. Age groups, 41-50 (HR = 1.38, CI = 1.25, 1.52) and Over 50 (HR = 1.99, CI = 1.76, 2.45) also had higher HRs than 31-40, those over 50 had higher ratios than those 41-50 (HR = 1.44, CI = 1.28, 1.62). Thus, younger clients tended to have longer survival time than older clients.

TABLE 3: SURVIVAL ANALYSIS RESULTS FOR ADAP-TO-VITAL STATS MATCH: TIME TO DEATH SINCE HIV DIAGNOSIS

Variable	Group	Param Estimate	p-value	HR	95% CIs	
Gender	Male	***	***	***	***	***
	Female	-0.04	.585	0.96	0.82	1.12
Race/ Ethnicity	White	***	***	***	***	***
	African American	0.12	.044	1.13	1.00	1.26
	Hispanic/Latino	0.18	.001	1.20	1.08	1.34
	Other	0.01	.941	1.01	0.72	1.42
Age Group	18-30 years old	***	***	***	***	***
	31-40 years old	0.13	.165	1.13	0.95	1.36
	41-50 years old	0.45	.000	1.57	1.31	1.87
	Over 50 years old	0.81	.000	2.26	1.86	2.74
Location	***	-0.03	.697	0.97	0.83	1.13
AIDS	***	1.69	.000	5.40	4.83	6.03
Insurance	***	-0.19	.000	0.83	0.75	0.90
Access	***	-0.07	.000	0.94	0.93	0.94

ADAP-to-HARS Survival Analysis

AIDS-Diagnosed ADAP Population. Table 4 shows the demographic characteristics for the 13,224 AIDS-diagnosed ADAP clients. This subset of ADAP clients was highly similar to the overall on gender and race/ethnicity population. Ninety percent were males and ten percent were females. Whites were 45 percent of AIDS-diagnosed clients, African Americans were 18 percent, and Hispanic/Latinos were 30 percent of this group. A subtle difference between AIDS-diagnosed clients and the overall population occurred in age group and geographical location. There were fewer AIDS-diagnosed clients between 18-30 years old (11 percent), a comparable amount between 31-40 years old (45 percent), and slightly more AIDS-diagnosed clients in the older age groups (32 percent between 41-50 years old and 12 percent Over 50 years old). There were slightly fewer AIDS-diagnosed clients residing in Title I/urban areas (90 percent) than the overall population, but slightly more clients residing in Title II/rural areas (nine percent).

TABLE 4: DEMOGRAPHICS FOR AIDS-DX ADAP CLIENTS

GENDER	FREQ	PCT
Male	11,799	89.22%
Female	1,425	10.78%
Other/Unknown	0	0.00%
RACE/ETHNICITY	**FREQ**	**PCT**
White	5,908	44.68%
African American	2,441	18.46%
Hispanic/Latino	4,010	30.32%
Other/Unknown	865	6.54%
AGE GROUP	**FREQ**	**PCT**
Under 18 years old	4	0.00%
18-30 years old	1,513	11.44%
31-40 years old	5,904	44.65%
41-50 years old	4,199	31.75%
Over 50 years old	1,583	11.97%
Unknown	21	0.00%
GEOGRAPHICAL LOCATION	**FREQ**	**PCT**
Title I/Urban Areas	11,967	90.49%
Title II/Rural Areas	1,200	9.07%
Unknown	57	0.00%
TOTAL	**13,224**	**100.00%**

The remaining variables of interest are shown in Table 5. Unlike the overall population in the insurance coverage group, fewer AIDS-diagnosed ADAP clients were ADAP Only (61 percent) and more had ADAP with Medicaid or private insurance benefits (39 percent). This difference may reflect the ability of less healthy clients to access the Medicaid system.

The most common mode of exposure or risk category, a variable available through the HARS match and not in the overall ADAP population, was MSM (64 percent), followed by IDU (13 percent), and MSM and IDU (nine percent). Heterosexual contact accounted for nine percent of all AIDS-diagnosed ADAP clients, and the remaining six percent of clients fell into the other category.

In comparison to HARS surveillance data in the same time period (December 2000), fewer AIDS-diagnosed MSM were present in ADAP than in HARS (76 percent), more IDU and heterosexual contact were in ADAP than in HARS (eight percent and two percent, respectively), and a similar percentage of MSM and IDU appeared in both groups (10 percent in HARS). These findings imply that ADAP is serving a fairly different AIDS subpopulation than the overall AIDS population by mode of exposure.

TABLE 5: MODE OF EXPOSURE FOR AIDS-DX ADAP CLIENTS

COVERAGE GROUP	FREQ	PCT
ADAP Only	8,093	61.20%
Not ADAP Only	5,131	38.80%
MODE OF EXPOSURE	**FREQ**	**PCT**
MSM	8,434	63.78%
IDU	1,688	12.76%
MSM and IDU	1,190	9.00%
Heterosexual contact	1,152	8.71%
Other	760	5.75%
TOTAL	**13,224**	**100.00%**

For Coverage Group, Not ADAP Only includes ADAP with Medicaid or private insurance benefits.

OA estimated survival curves and performed survival analyses on the ADAP-to-HARS match in the same manner OA analyzed the ADAP-to-Vital Statistics match.

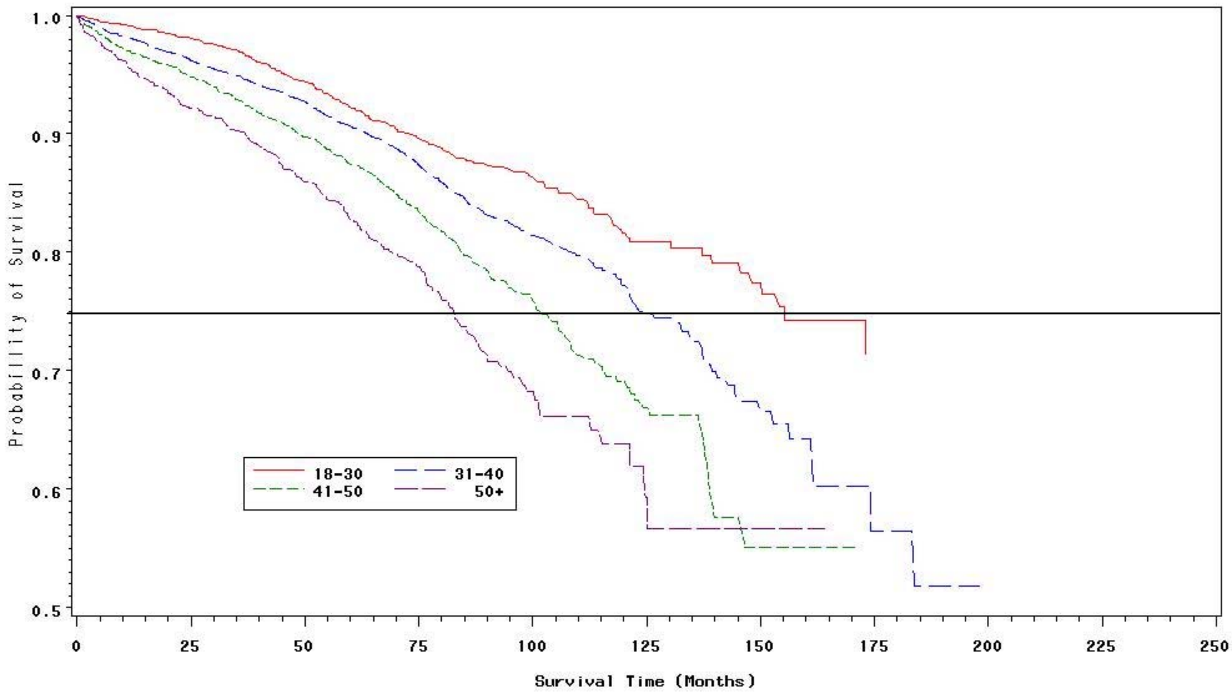
<u>Survival Curves</u>. (Kaplan-Meier estimates of the 25 percentile of survival time since AIDS diagnosis, or time in months when 25 percent of the AIDS clients died, was examined for the variables of interest). Reliable estimates for the 50 and 75 percentile were not available due to the low number of deaths. The survival curves by age group, mode of exposure, and insurance coverage group are shown in the following figures.[2]

Figure 5 shows the survival curves by age group. OA found that the 25 percentile of survival time declined as age increased. That is, AIDS-diagnosed 18-30 year olds had the longest survival times (155.21 months), followed by 31-40 year olds (125.71 months), 41-50 year olds (101.80 months), and Over 50 year olds (82.89 months). The log-rank test statistic was significant (<u>p</u><.000) indicating group differences across age groups, without adjusting for other variables.

[2] As with ADAP-to-Vital Statistics analysis, only survival curves for categorical variables with a significant long-rank test are shown.

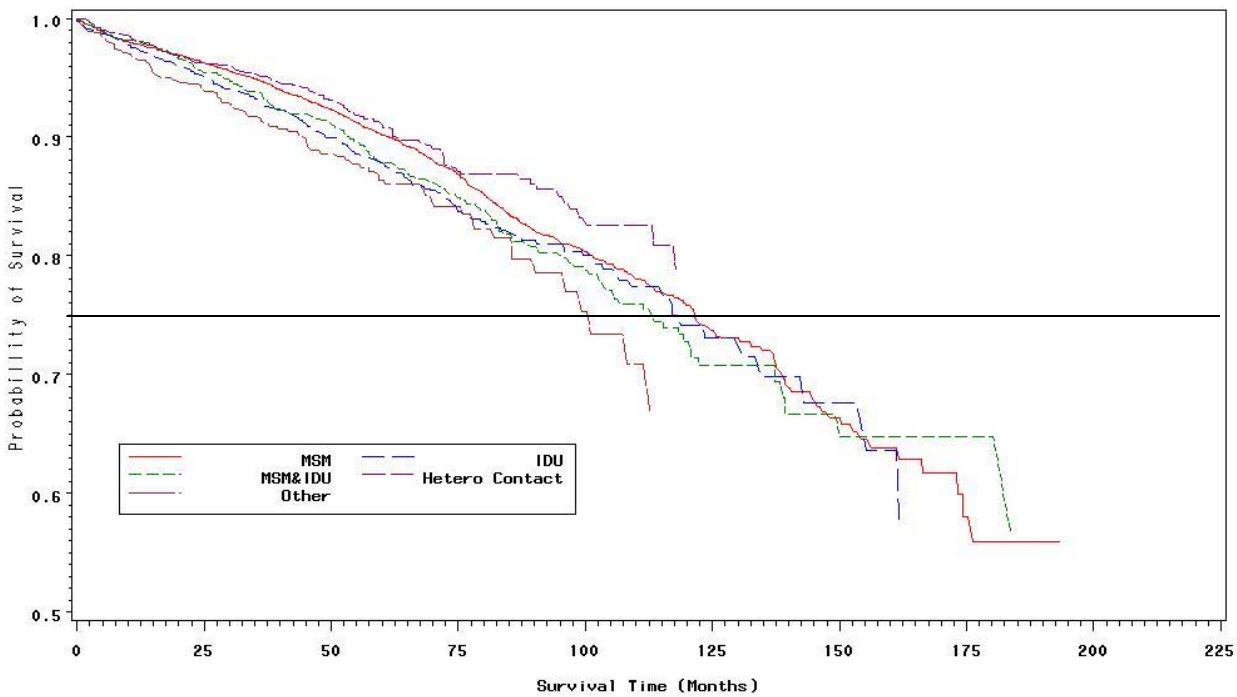
Survival patterns by mode of exposure are shown in Figure 6. Longer 25 percentile of survival time estimates were found among MSM (121.38 months) and IDU (118.79 months) than MSM and IDU (113.05 months) or Other category (101.12 months). A reliable estimate for AIDS-diagnosed clients via heterosexual contact could not be computed. The log-rank test statistic was significant (p<.002).

Figure 5. Kaplan-Meier Estimate of Survival Curves by Age Group for AIDS-Diagnosed Clients



Department of Health Services, Office of AIDS, AIDS Drug Assistance Program

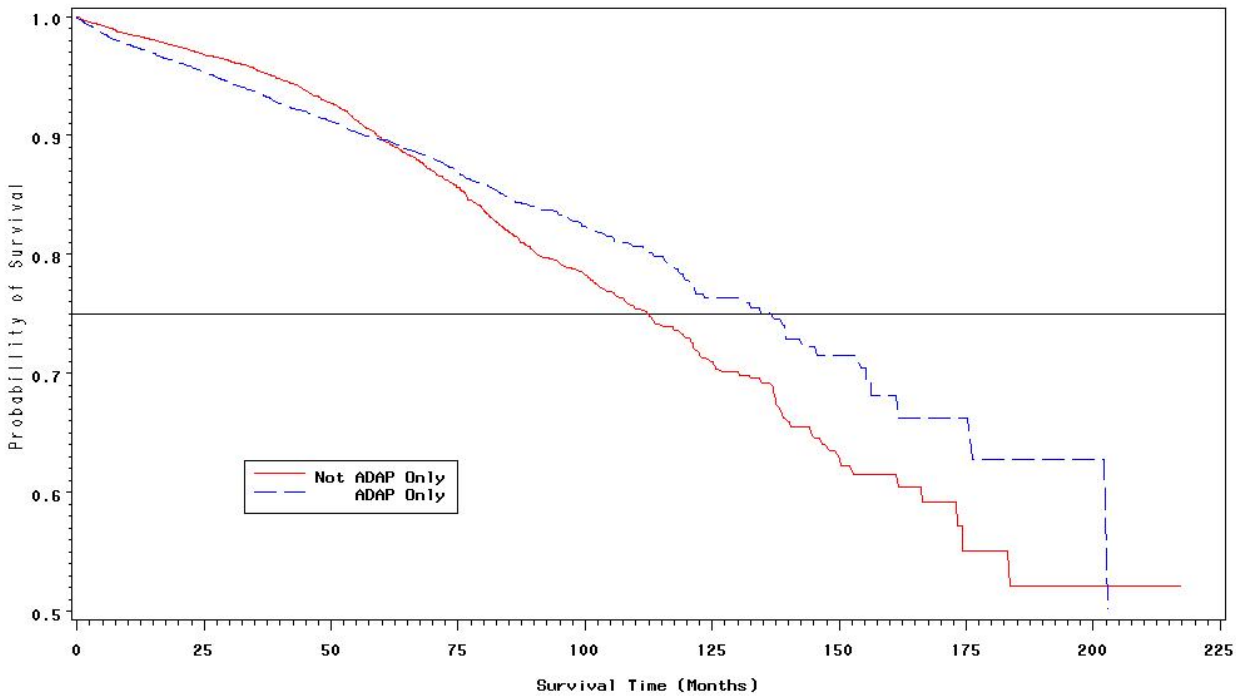
Figure 6. Kaplan-Meier Estimate of Survival Curves by Mode of Exposure for AIDS-Diagnosed Clients



Department of Health Services, Office of AIDS, AIDS Drug Assistance Program

Figure 7 shows the survival curves for AIDS-diagnosed clients by insurance coverage group. Similar to the ADAP-to-Vital Statistics match, ADAP Only clients (137.18 months) had longer 25 percentile of survival time estimates than Not ADAP Only clients, or those with Medicaid or private insurance benefits (112.20 months). The log-rank statistic was significant (p<.018).

Figure 7. Kaplan-Meier Estimate of Survival Curves by Insurance Coverage Group for AIDS-Diagnosed Clients



Department of Health Services, Office of AIDS, AIDS Drug Assistance Program

<u>Time to Death Since AIDS Diagnosis</u>. Of the 13,224 AIDS-diagnosed ADAP clients, there were 1,797 deaths for a case fatality rate of 14 percent. This was much higher than the eight percent rate for the overall HIV/AIDS ADAP population.

Table 6 shows the results of the survival analysis from AIDS diagnosis date-to-death using a Cox-Proportional Hazard Model. Comparable to the ADAP-to-Vital Statistics match, significant differences in death hazards were found on race/ethnicity, age group, insurance coverage group, and access months. In contrast, the specific group differences within variables differed between populations on race/ethnicity and age group. AIDS-diagnosed African Americans had a significantly lower HR than Whites (HR = 0.81, CI = 0.71, 0.92) indicating longer survival time for African Americans. Higher HRs were found between 31-40 year olds (HR = 1.45, CI = 1.25, 1.67), 41-50 year olds (HR = 2.05, CI = 1.76, 2.39), and Over 50 (HR = 2.74, CI = 2.28, 3.30) year olds than 18-30 year olds. This indicated longer survival time for the youngest age group.

Comparable within group differences were found on insurance group coverage and access months for both AIDS-diagnosed clients and the overall ADAP population. ADAP Only clients had smaller HRs or longer survival time than those with Not ADAP Only clients (HR = 0.82, CI = 0.74, 0.91), and clients with more access months had smaller HRs or longer survival time than others (HR = 0.94, CI = 0.93, 0.94).

Additional models were performed to test all pairwise comparisons between other groups for race/ethnicity, age group, and mode of exposure. Hispanic/Latinos had a higher HR or shorter survival time than African Americans (HR = 1.36, CI = 1.17, 1.56). Age groups, Over 50 (HR = 1.90, CI = 1.63, 2.21) and 41-50 (HR = 1.42, CI = 1.27, 1.59) had higher HRs or shorter survival times than 31-40 year olds. Over 50 years also had a higher HR than 41-50 year olds (HR = 1.33, CI = 1.14, 1.57). Among mode of exposure, only one significant difference was found. The Other category had a higher HR than IDU (HR = 1.36,
CI = 1.07, 1.74).

TABLE 6: SURVIVAL ANALYSIS RESULTS FOR ADAP-TO-HARS MATCH: TIME TO DEATH SINCE AIDS DIAGNOSIS

Variable	Group	Param Estimate	p-value	Hazard Ratio	95% Confident Intervals	
Gender	Male	***	***	***	***	***
	Female	-0.19	.092	0.83	0.66	1.03
Race/ Ethnicity	White	***	***	***	***	***
	African American	-0.21	.001	0.81	0.71	0.92
	Hispanic/Latino	0.09	.127	1.09	0.98	1.23
	Other	-0.22	.188	0.81	0.59	1.11
Age Group	18-30 years old	***	***	***	***	***
	31-40 years old	0.37	.000	1.45	1.25	1.67
	41-50 years old	0.72	.000	2.05	1.76	2.39
	Over 50 years old	1.01	.000	2.74	2.28	3.30
Mode of Exposure	MSM	***	***	***	***	***
	IDU	-0.10	.207	0.90	0.77	1.06
	MSM and IDU	-0.05	.507	0.95	0.81	1.11
	Heterosexual contact	-0.08	.544	0.93	0.73	1.19
	Other	0.21	.063	1.23	0.99	1.53
Location	Title I/Urban Areas	***	***	***	***	***
	Title II/Rural Areas	-0.16	.079	0.86	0.72	1.02
Insurance	ADAP Only	***	***	***	***	***
	Not ADAP Only	-0.20	.000	0.82	0.74	0.91
Access	***	-0.07	.000	0.94	0.93	0.94

DISCUSSION

This study examined survival rates in ADAP's overall PLWH/A population and its AIDS-diagnosed subpopulation. In the overall population, African Americans and Hispanic/Latinos had shorter survival times since HIV diagnosis than Whites. This finding suggests that people of color may delay seeking treatment and that people of color may encounter socioeconomic and cultural barriers to educational and preventive services, primary care and treatment, and access to costly HAART drug therapies. These barriers include, but are not limited to, stigma, distrust of providers, poor health literacy, poverty, and homelessness (Morin, Kahn, Richards, and Palacio, 2000; Kalichman and Rompa, 2000; Russell, 2002; Wong and Xing, 2003). In the AIDS-diagnosed subpopulation, in contrast, Whites and Hispanic/Latinos had shorter survival times than African Americans. Because of the fewer number of cases among the AIDS-diagnosed group than the overall PLWH/A group, cases and deaths, these results were less stable and more difficult to interpret.

Older age groups tended to have shorter survival times than younger age groups in both PLWH/A and AIDS-diagnosed groups, and for PLWH/A only, AIDS-diagnosed clients had shorter survival times than their HIV-only counterparts. These findings were not surprising since the progression from HIV infection to AIDS may take up to ten years or longer to develop. AIDS, by definition, is a more serious condition than HIV because one is highly susceptible to diseases that a healthier person can resist.

In both groups, ADAP Only clients had longer survival times than those with Medicaid or private insurance benefits. It was unclear as to why this occurred. ADAP Only provides prescription drug services whereas Medicaid and private insurance coverage includes both drug and medical services. Medicaid clients may be disabled and sicker than ADAP Only clients. Because OA does not have access to other drugs not paid for by ADAP for clients with Medicaid or private insurance, the effects of various HAART therapies could not be examined. Also, clients with more access months had longer survival times than those with fewer access months. This may also be related to ADAP clients transitioning to Medicaid because of poorer health.

Mode of exposure was examined for the first time in ADAP for AIDS-diagnosed clients only. ADAP had a smaller percentage of AIDS-diagnosed MSM but a larger percentage of IDU and heterosexual contact than in the HARS population. A similar percentage of MSM and IDU occurred in both groups. Again, these findings suggest ADAP enrolls and serves a fairly different AIDS subpopulation than the overall AIDS population (by mode of exposure, or risk category).

A few limitations to this study should be noted. First, different levels of completeness in the matching identifiers prevented a perfect match between ADAP, Vital Statistics, and HARS data sets. For example, there may be some misclassifications in the

ADAP-to-HARS match but the matching criterion was setup to maximize true matches and minimize false matches. Second, the low case fatality rate (deaths/cases) particularly in the AIDS-diagnosed group only allowed reliable estimates for 25 percentile of survival time. The 50 percentile (median) and 75 percentile of survival times could not be computed for many groups. Third, more complete insurance coverage data were not available for a more in-depth comparison of ADAP Only versus Not ADAP Only clients or those with public versus private insurance. Fourth, the results for California's ADAP may not be representative of other state ADAPs or to the United States population of PLWH/A.

California's ADAP will continue to monitor the effectiveness of the program by tracking clients' health indicators such as CD4 counts and viral load (Wong and Fairgrieves, 2004) and health outcomes such as morbidity and mortality rates (Wong and Xing, 2003). These studies will give health care providers, practitioners, and advocates a better idea of where to focus their efforts to provide the best HIV/AIDS services based on the different socioeconomic levels and cultural values of the population it serves and to try to minimize the existing barriers to education, prevention, and care and treatment services.

REFERENCES

Bhattacharya, J., Goldman, D., Sood, N. The link between public and private insurance and HIV-related mortality. *Journal of Health Economics*. 2003; 22: 1105-1122.

Bryan, J., Sun, R. Temporal trends in survival after AIDS in California, 1985-1994. *California HIV/AIDS Update* 1998 April; 11: 28-29.

California Department of Health Services, *Updated graphs on temporal trends in survival after AIDS in California,* Sacramento, CA (updated 2002 May 01). Available at: http://www.dhs.ca.gov/aids/Statistics/Graphs/pps/mar2002.pps.

Department of Health and Human Services (DHHS) and Henry J. Kaiser Family Foundation, *Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents,* 1999 May 5.

Fellegi, I.P., Sunter, A.B. A theory of record linkage. *Journal of the American Statistical Association.* 1969; 64: 1183-1320.

Kalichman, S.C., Rompa, D. Functional health literacy is associated with health status and health-related knowledge in people living with HIV-AIDS. *J Acquir Immune Defic Syndr.* 2000;25:337-344.

Lee, L.M., Karon, J.M., Selik, R., Neal, J.J., Fleming, P.L. Survival after AIDS diagnosis in adolescents and adults during the treatment era, United States, 1984-1997. *JAMA.* 2001; 285: 1308-1315.

Morin, S.F., Kahn, J.G., Richards, T.A., Palacio, H. *Eliminating racial and ethnic disparities in HIV care: The California report.* San Francisco (CA): AIDS Policy Research Center & Institute for Health Policy Studies and AIDS Research Institute, University of California, San Francisco, Policy Monograph Series--2000 Mar.

Office of AIDS. *Acquired Immunodeficiency Syndrome (AIDS) HIV-AIDS Reporting System Surveillance Report for California – December 31, 2000.* Available at http://www.dhs.ca.gov/aids/Statistics/default.htm.

Rogot, E., Sorlie, P., Johnson, N.J. Probabilistic methods in matching census samples to the National Death Index. *Journal of Chronic Diseases*. 1986; 39: 719-734.

Russell, S. Blacks dying of AIDS faster: Many in San Francisco slow to seek care. San Francisco Chronicle. 2002 Jun 27;Sect. A:15

Wong, D.T., Fairgrieves, K.S. *CD4 Counts and Viral Load Measurements in AIDS Drug Assistance Program (ADAP): Quality (Data) Management and Health Indicators*. California Department of Health Services, Office of AIDS, 2004.

Wong, D.T., Xing, Biao. *Morbidity and mortality rates in California's AIDS Drug Assistance Program.* California Department of Health Services, Office of AIDS; 2002.

APPENDIX A

INSTRUCTIONS FOR SOUNDEXING

The purpose of soundexing is to facilitate matching and unduplicating reported HIV and AIDS cases. The soundex code maintains the confidentiality of reported cases by converting the last name of an individual to an index letter and a three-digit number. In coding by this system, the index letter is the first letter of the last name and the subsequent letters are converted to a numeric code in accordance with the following general rules:

Rule	Instructions	Example
1	The first letter of the last name is never coded.	
2	The vowels A, E, I, O, U, and Y are never coded.	
3	The consonants H, and W are never coded.	
4	Key letters and their equivalents are converted to code numbers. Key Letter — Equivalents — Code Number B — B,F,P,V — 1 C — C,G,J,K,Q,S,X,Z — 2 D — D,T — 3 L — L — 4 M — M,N — 5 R — R — 6	
5	The consonants of the last name, other than the first letter and H and W, are converted to their respective code numbers in the order in which they appear in the name.	HOLMES H452 45 2 GWILFOYLE G414 41 4
6	The numeric code always consists of three digits. The codes for names which do not contain three key-letters or their equivalents are completed by adding zeros. Note that the zeros follow the assigned number code.	GRAHAM G650 6 5 BAILEY B400 4 SHAW S000
7	The soundex code for names that contain more than three key-letters, or their equivalents, are complete when a three-digit numeric code has been assigned.	VONDERLEHR V536 53 6 4 6
8	Two or more key-letters, or their equivalents, appearing together are treated as one key-letter and are assigned one number.	BALLOU B400 4 - JACKSON J250 2 - - 5
9	A key letter, or its equivalent, immediately following an initial letter (first letter of the last name) of the same group or value is not coded.	SCANLON S545 - 5 4 5 SCKLAR S460 - - 4 6

APPENDIX A — CONTINUED

Rule	Instructions	Example
10	Key-letters, or their equivalents, separated by A,E,I,O,U, or Y are coded separately.	HANNON H550 5 - 5 SALKIEWICS S422 42 2 -
11	Key-letters, or their equivalents, separated only by the letter W or the letter H are coded as one key-letter. Note that in the name Schkolnik, the C is not coded because it is in the group equivalent to the letter S, and the first K is not coded because it is in the group equivalent to the letter C, from which it is separated only by an H.	SOKWZY S200 2 - SCHKOLNIK S452 - - 45 2
12	Abbreviated prefixes such as Mc or St. are coded as if spelled out.	MCKILHAN=MACKILHAN M245 2 - 4 5 ST. JOHN=SAINT JOHN S532 53 2 -
13	An apostrophe in a name is disregarded.	O'NEILL O540

APPENDIX B

MATCHING METHOD

1. Introduction

Matching two independent data files is a record linkage process in which records from the two files are linked based on the agreement of a set of identifiers (e.g., names, SSN, gender, DOB). A common problem with this type of matching is that there are always false matches and false non-matches due to non-uniqueness, incomplete information, data entry, and coding errors in the identifiers.

Successful matching depends on a suitable assessment of the matches and non-matches regarding whether they are true or false. In this respect, the matching problem is truly a classification problem. To produce acceptable matching, OA wanted to minimize the false match and false non-match rate. Or equivalently, OA wanted to maximize the sensitivity and specificity of the classification.

Based on a record linkage theory by Fellegi and Sunter (1969), Rogot, Sorlie, and Johnson (1986), developed a practical probabilistic method for matching census samples to the National Death Index. The basic idea of their method is to compute a weight for each identifier used in matching based on its probability of occurrence and then construct a score for each match using those weights. A suitable cut-off is then chosen for the classification. The matching criteria are assumed to be conservative so that false non-matches are expected to be small and can be disregarded.

In this study, OA matched ADAP clients with Vital Statistics death files and HARS cases. From 1998-2000, ADAP data had 30,834 records (including duplicates in SSN); Vital Statistics had 926,945 death records in California between 1998-2001; and HARS data had 126,269 records as of July 31, 2002.

2. Matching Method

Using the same idea as Rogot, Sorlie, and Johnson (1986), OA developed a probabilistic approach for our matching and classification described below.

2.1 Matching Criteria

Possible identifiers used for matching included SSN, names (first and/or last name), soundex (for records where last names were not available), DOB, gender, and race/ethnicity. Multiple criteria were considered for cases with incompleteness on matching identifiers.

APPENDIX B — CONTINUED

For matching ADAP-to-Vital Statistics data, OA used the following matching criteria: (1) SSN; or (2) last name + DOB + gender. For matching ADAP-to-HARS data, OA used the following criteria: (1) SSN; or (2) soundex + DOB + gender. First name and race/ethnicity was not used due to coding differences between data sets. For example, a shortened name (Mike or Tom) may be recorded in one source, but the full name is recorded in the other (Michael or Thomas).

Data records from the two different sources that agreed on either or both of these criteria were called potential matches, and those records that failed to agree on any of these criteria were called non-matches. OA believed this criteria was conservative enough such that the false non-match rate was assumed to be small and not assessed. OA was, however, concerned with the potential matches, which usually consist of some number of false matches. Thus, OA wanted to reclassify the potential matches into true and false matches in a way that was objective, consistent, and tractable.

2.2 Scoring and Reclassifying Potential Matches

All potential matches were classified into three mutually exclusive classes.

Class 1: SSN matched and at least two of the following three criteria matched—soundex/last name, DOB, or gender.

Class 2: All other situations different from that of Class 1 and 3.

Class 3: (1) SSN matched, but:
- (1-a) Soundex/Last name existed but did not match
- (1-b) First name existed but did not match
- (1-c) DOB existed but did not match.

(2) Soundex/Last name, DOB, and gender all matched, but:
- (2-a) SSN existed but did not match
- (2-b) First name existed but did not match.

Class 1 was labeled as <u>confirmed true matches</u>, and Class 3 was labeled as <u>confirmed false matches</u>. Class 2 was subject to reclassification for true or false matches. Since OA did not have sufficient information to do so deterministically, OA used a probabilistic approach to do the re-classification. Class 4 was non-matches so no classification was needed here.

A probabilistic scoring method for assessing and reclassifying potential matches was developed as follows. OA first calculated a weight for each identifier that was used in matching and for an additional assessment. The weight was defined as

APPENDIX B — CONTINUED

the logarithm (with base 2) of the inversed probability that an attribute of an identifier appearing in the target data source (e.g., Vital Statistics or HARS). For example, if the last name

"Johnson" had an occurring probability of 0.0002, then its weight was $\log_2(1/0.0002) \approx 12.2877$. Then, OA used the weights to construct a score for each potential match. Each score was defined as:

Score = W_SSN * IND(SSN matched)
+ W_DOB * IND(DOB matched)
+ W_gender * IND(gender matched)
+ W_last_name * IND(last name matched)
+ W_first_name * IND(first name matched)

for ADAP-to-Vital Statistics data matching, and;

Score = W_SSN * IND(SSN matched)
+ W_soundex * IND(soundex matched)
+ W_DOB * IND(DOB matched)
+ W_gender * IND(gender matched)
+ W_last_name * IND(last name matched)
+ W_first_name * IND(first name matched)
+ W_race/ethnicity.* IND(race/ethnicity matched)

for ADAP-to-HARS data matching, where indicator function IND(…) = 1 if its content was true and 0 otherwise.

Race/Ethnicity was not used in the ADAP-to-Vital Statistics match because of different coding schemes. The higher the score a potential match had, the greater the likelihood it was a true match. Each score was used to screen false matches. It was also used to eliminate duplicates of records caused by multiple matching (i.e., one ADAP records matching with several Vital Statistics or HARS records or vice versa).

2.3 Validation and Cut-Off Selection

There may be other classification rules that were applicable to our matching procedure. OA choose the above rules, because OA wanted to draw samples from the confirmed true (Class 1) and false-matches class (Class 3) and use them to select a suitable cut-off value from the weighted scores to reclassify Class 2. Ideally, OA wanted to have independent validation samples that contained known true and false matches to the target data. Then, OA could have used them to

APPENDIX B— CONTINUED

select the cut-off point. Since no independent validation samples were available, OA drew internal data samples from confirmed true and false matches to construct validation samples.

Using the validation sample, OA selected a cut-off value that maximized the product of sensitivity and specificity of the classification. Here sensitivity is defined as the probability that a true match was classified as a "true match," and specificity as the probability that a false match was classified as a "false match." For the validation sample for the ADAP-to-Vital Statistics matching, a cut-off of 20.9 gave the optimal sensitivity of 100 percent and specificity of 100 percent. For the validation sample for ADAP-to-HARS matching, a cut-off value of 27.7 gave the optimal sensitivity of 99 percent and specificity of 100 percent. The sensitivity and specificity plots for the two validation samples are shown in Figure A.

Next, OA applied the cut-off value to reclassify Class 2. If a potential match had a score higher than the cut-off value, then it was classified as a true match. Otherwise, it was classified as a false match.

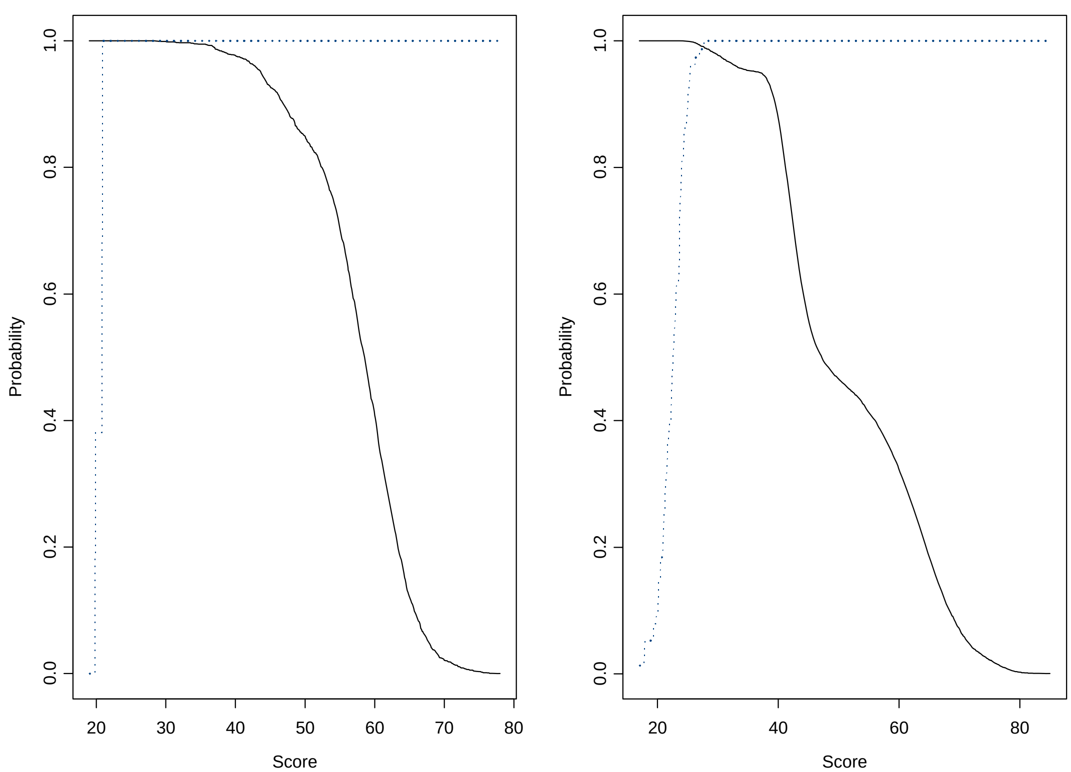


Figure A. Sensitivity and Specificity of Classification. Left: Validation sample for ADAP-to-Vital Statistics matching. Right: Validation sample for ADAP-to-HARS matching. Solid line: Sensitivity. Dotted line: Specificity.

APPENDIX B — CONTINUED

3. Matching Results Summary

3.1 ADAP-to-Vital Statistics Matching

As described above, OA first performed data matching on SSN. Then OA proceeded to match soundex + DOB + gender. After removing duplicates and multiple matches by keeping the one with highest scores, OA obtained a total of 30,309 unique ADAP records of which there were 2,341 potential matches and 27,968 non-matches.

The 2,341 potential matches were classified into three classes using the rules as described above (see Table 3.1). Class 1 and 3 were regarded as true matches and false matches, respectively. Class 2 consisted of potential matches that could not be directly determined due to insufficient information on the identifiers. A cut-off value of scores was selected based on Class 1 and 3 as described above and then applied to Class 2 for a probabilistic reclassification into true or false matches. A cut-off value of 20.9 was chosen, which yielded 2,320 true matches, 21 false matches, and 27,968 non-matches. In sum, there were 2,320 matches and 27,989 non-matches.

Table 3.1. Classification of ADAP-to-Vital Statistics Matching Results

Class	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1,861	6.14%	1,861	6.14%
2	459	1.51%	2,320	7.65%
3	21	0.07%	2,341	7.72%
4	27,968	92.28%	30,309	100.00%

Note: Class 1 was confirmed true matches. Class 2 was to be determined. Class 3 was confirmed false matches. Class 4 were non-matches.

3.2 ADAP-to-HARS Matching

OA used the exact same procedure as in the ADAP-to-Vital Statistics match described above. Of the 30,309 ADAP records, there were 17,825 potential matches and 12,484 non-matches. After classifying the 17,825 potential matches into true (Class 1) and false (Class 3) matches, OA applied a cut-off value of 27.7 for those without sufficient information for proper classification. The end result yielded 13,224 true matches, 4,601 false matches, and 12,484 non-matches or 13,224 matches, and 17,085 non-matches.[2]

[2] Sixty-three matches were reclassified as non-matches, because the persons had died prior to 1998.

APPENDIX B — CONTINUED

Table 3.2. Classification of ADAP-to-HARS Matching Results

Class	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	8,949	29.53%	8,949	29.53%
2	8,800	29.03%	17,749	58.56%
3	76	0.25%	17,825	58.81%
4	12,484	41.19%	30,309	100.00%

Note: Class 1 was confirmed true matches. Class 2 was to be determined.
Class 3 was confirmed false matches. Class 4 were non-matches.